1939

A performance test for school age children wtih a language handicap

Jennylouise Lockwood White

WASHINGTON UNIVERSITY

Central Institute for the Deaf

A PERFORMANCE TEST FOR SCHOOL AGE CHILDREN WITH A LANGUAGE HANDICAP

by

Jennylouise Lockwood White

A dissertation presented to the Board of Graduate Studies of Washington University in partial fulfilment of the requirements for the degree of Master of Science in Education

June, 1939

Saint Louis, Missouri

TABLE OF CONTENTS

CHAPTERS

APPENDICES

TABLES

FIGURES

CHAPTER I.

INTRODUCTION

The rapid growth of democracy in education in recent years has brought formal schooling within the reach of many who would have been denied this opportunity a century ago. In order for education to be efficient and economical, as well as widespread, educators have begun to recognize the necessity for some type of measurement of the individual child's capacities, so that he may be trained in the manner best suited to his ability as well as to the needs of society. Thus far, the most satisfactory measurements developed are intelligence tests, of which there are a great variety.

However, these tests, most of which involve the use of language, are totally inadequate to measure the native ability of the deaf child with his great language handicap. The demand for an instrument designed to suit his individual needs has given rise to investigations, resulting in the production of several non-language scales, but so far, none comparable to the best of the linguistic scales have been devised. If a test is to be an accurate estimate of the deaf child's intelligence, it must be constructed by some one who is acquainted not only with methods of intelligence testing in general and of performance testing in particular, but with the deaf child himself.

The question of whether or not the intelligence of the deaf child is equal to that of the hearing child, other factors being equal, has also been studied by many investigators, but it is still a controversial issue. However, before this

problem can be solved, it is necessary to define intelligence in a manner that will be acceptable to most educators.

Definitions of intelligence are many and varied, ranging from Terman's viewpoint, that it is the ability "to carry on abstract thinking,"[1] to that of Colvin, that it is the individual's ability "to adjust himself to his environment."[2] If the former is accepted, one must infer that most adults do not use intelligence often, since they seldom think in purely abstract terms, and the average child of preschool age, as well as all deaf mutes, would have no intelligence. On the other hand, to accept the latter definition would be to admit that the lower forms of animal life, and even plants, have intelligence.[3] Most psychologists have accepted definitions somewhere between these two, using the judgment of the world as the best criterion. It is generally agreed that intelligence should include not only adaptation to environment, but the ability of the individual

> ..... to reconstruct the factors of his environment in accordance with the most fundamental needs of himself and his group.[4]

According to Pintner, most psychologists have accepted Stern's definition of intelligence, which states

> Intelligence is a general capacity of an individual consciously to

---

1. Symposium, "Intelligence and Its Measurement," Journal of Educational Psychology, 12: 123, 195, 1921.
2. Ibid., loc. cit.
3. Boynton, Paul L., Intelligence, New York: Appleton and Co., 1933, pp. 10-11.
4. Ibid., pp. 19-20.

> adjust his thinking to new requirements; it is general mental adaptability to new problems and conditions of life.[5]

Many psychologists differentiate between the types of intelligence. Pintner calls them "verbal, concrete, and social;" Bridges, "cognitive, conative, and affective;" and Thorndike, "abstract, mechanical, and social."[6] Spearman distinguishes between "general and specific factors" in intelligence,[7] while Thorndike believes that

> a number of specific abilities exist which are highly intercorrelated because of the elements they possess in common.[8]

Binet expresses this opinion in saying that one must test the "complex processes" to get a true measure of an individual's ability.[9]

Terman, Kenmon and others believe that "conceptual thinging" is of a higher order than "manipulative skill," but Trabue claims that it may not be.[10] Drever and Collins say that in the past, abstract intelligence has been regarded as the higher type, and intelligence tests have been made to measure it mainly, since the criterion in selecting the tests has been the extent to which the results correlate with school and teachers' estimates, both greatly weighted

---

5. Pintner, Rudolf and Paterson, Donald G., A Scale of Performance Tests, New York: Appleton and Co., 1931, p. 2.
6. Schieffelin, Barbara and Schwesinger, Gladys C., Mental Tests and Heredity, New York: Galton Publishing Co., 1930, p. 13.
7. Boynton, op. cit., pp. 22-25.
8. Schieffelin and Schwesinger, op. cit., p. 14.
9. Ibid., p. 4.
10. Ibid., loc. cit., and p. 260.

on the side of abstract intelligence. They do not test adequately the ability to deal with various situations, nor the handling of concrete material, which are both necessary in a child's education. These authors believe that language tests must be supplemented by performance tests, if they are to test the "complex processes" of normal individuals.[11] Pintner also stresses the complexity of intelligence, and states that one kind of test will not measure it adequately.[12] Wells does not believe that a performance test measures the same thing as a verbal test, but believes that a true measurement requires all types of tests.[13] Binet included some performance items, but his scale is mainly linguistic.[14]

Experimental evidence shows a fairly high positive correlation between linguistic and performance tests, indicating that there must be a common factor in the performance called for on both types of tests.[15] Occasionally, however, a child does well on one test, and poorly on the other. In such cases, it is important to know this fact, and educate the child accordingly, so the performance test acts as a check on the linguistic test, or shows the unequal development between verbal and non-verbal abilities.[16]

---

11. Drever, James and Collins, Mary, Performance Tests of Intelligence, Edinburgh: Oliver and Boyd, 1928, pp. 9-10.
12. Pintner and Peterson, op. cit., p. 21.
13. Wells, F. L., Mental Tests in Clinical Practice, New York: World Book Co., 1927, p. 315.
14. Drever and Collins, loc. cit.
15. Schieffelin and Schwesinger, op. cit., pp. 66-68.
16. Arthur, Grace, A Point Scale of Performance Tests, New York: The Commonwealth Fund, Division of Publications, 1930, 1:5-9.

In addition, the performance test has proved itself valuable as a supplement to the Binet in establishing self-confidence in pupils before giving the Binet, and a better attitude in cases of emotional repression, as well as in providing the examiner an excellent opportunity to observe the individual's behavior.[17] One of its chief advantages is its universal appeal to the subject.[18] Another advantage is that training, either at home or at school, does not have as great an effect on the results of a performance test as it does on the results of a language test.[19]

Although its value may be great in supplementing the language test on the normal individual, the performance test has an even greater value: that of testing the mental ability of individuals who, because of a language handicap, cannot be measured adequately by a linguistic test. This group includes illiterates, foreigners, and cases with a speech disability, as well as those with such a physical abnormality as aphasia or deafness.[20] The question immediately arises as to whether the performance test can take the place of the linguistic test. If Terman's definition of intelligence is accepted, it must be admitted that performance tests cannot test the higher levels of intelligence, and that an adequate measure of the intelligence of the deaf and other linguistic-

17. Arthur, op. cit., pp. 8-12.
18. Schieffelin and Schwesinger, op. cit., p. 69.
19. Drever and Collins, op. cit., p. 15.
20. Ibid., pp. 13-15.

ally-handicapped probably will never be found.[21] Dearborn, Shaw, Lincoln, Herring, and Squires believe that it is possible for performance test situations to call for analytic and synthetic thinking, but point out that the upper ranges of the present performance scales do not include tests difficult enough to call for abstract thinking.[22] However serious this defect may be, it refers to the contents of the tests, and not to their non-verbal nature, and may be remedied in time, when more is learned of the construction of tests.[23] In the meantime, since performance tests are the best instruments available for testing the deaf and others with a language handicap, it is the psychologist's task to strive continually to construct better ones, and to compare them with the most reliable of the language tests.

It is for this purpose that the present study has been made. While there are many good standardized performance tests available, comparatively few of these are suited to the needs of the deaf child of school age. Although many of them require no verbal response, they do require verbal directions, which involve lipreading and comprehension of language on the part of the subject, and are therefore not tests of intelligence alone, when used for the deaf.[24]

At Central Institute for the Deaf, where the study is

---

21. Schick, Helen F., "A Performance Test for Deaf Children of School Age," Volta Review, 36: 657, Nov. 1934.
22. Schieffelin and Schwesinger, loc. cit.
23. Lane, Helen Schick, "Measurement of the Mental and Educational Ability of the Deaf Child," Journal of Exceptional Children, 4:169, May, 1938.
24. Ibid., loc. cit. and p. 170.

being made, it is deemed advisable to give every child a retest annually, since school routine and discipline often make a marked difference in a pupil's habits of attention and concentration in a year's time, and therefore may alter his performance on the test.[25] Since one psychologist administers the entire testing program, the scale must be as economical of her time as possible, as well as short enough to avoid fatigue and loss of attention of the subject. It is generally agreed that behavior problems and attention necessitate an individual test for the deaf child. On the other hand, there must be as great a variety of tests as possible, so that all sides of the complex structure of intelligence will be measured.

Such a scale, fulfilling all these requirements, and composed of nine items from standardized performance scales, has been assembled by Lane, psychologist at Central Institute. Norms of the individual tests included are being used, but there is at present no standardization of the scale as a whole.[26] It is the purpose of this study to compute correlations between the scores of hearing children on this test, and mental ratings on standardized, reliable language tests, and the scores of deaf children on this test and on other performance tests, in order to determine the reliability of the scale. A comparison will be made between the hearing and deaf children on such a scale, and an attempt will be made to determine if

25. Lane, op. cit., p. 172.
26. Ibid., loc. cit.

the deaf as a group show superiority or retardation in specific tests.

CHAPTER II.

HISTORY

Although the measurement of intelligence is by no means a new idea, it was not until comparatively recent times that guesswork and subjective judgment were discarded in favor of more scientific methods.

The nineteenth century agitation for social justice for all classes, and the simultaneous growing scientific interest in the insane and feebleminded, gave rise to the movement for the education of the mentally retarded, begun by Itard and Seguin, and soon adopted by many European countries and the United States. The need for a satisfactory means of segregation of the backward child was apparent at once, but it was not until the latter part of the nineteenth century that psychologists turned their attention to this problem, which had long confronted the social reformer and educator.[27]

The first scientist to attempt an objective measurement of intelligence was Galton, whose study of eugenics necessitated a method of measuring individual differences. The anthropological studies of the nineteenth century also gave impetus to the movement, since it was necessary to have accurate measurements of intelligence for comparison with cephalic indices. At about the same time, the first attempts were being made to carry over the methods employed in experimental psychology, into the field of psychiatry.[28]

---

27. Pintner, Rudolf, Intelligence Testing. New York: Henry Holt and Co., 1923, pp. 3-13.
28. Ibid., pp. 13-23.

Oehrn is credited with being the first to make a scientific approach to intelligence testing, in 1889. The classification of her tests into perception, memory, association and motor ability showed her recognition of the complexity of intelligence.[29]

The term "mental test" was used for the first time in 1890 by Cattell, who initiated the testing movement in America. He advocated the standardization of methods of procedure and the establishment of norms, and in 1896, published the results of his tests on Columbia College students.[30] In this same year, Witmer of Pennsylvania University took a significant step by opening the first psychological clinic in this country.[31]

During this decade many others, in this country and abroad, were experimenting with tests of memory, imagery, imagination, attention, and motor ability, and were comparing the test results with other measures of mental ability, but in most cases the correlations were low. It was generally recognized that most of the tests used were lacking in reliability and validity, and tested only the lower mental processes. While this criticism tended to discourage investigation for a few years, it was during this period that the forerunner of the modern intelligence scale was devised. In 1902 Blin attempted to put definite questions in a graded scale. This was followed in 1905 by the first of Binet's

29. Boynton, op. cit., pp. 151-152.
30. Pintner, op. cit., pp. 14-15.
31. Boynton, op. cit., p. 156.

three scales, which introduced the mental age concept. While the first scale was intended only for dull individuals, his 1908 and 1911 revisions were useful for the classification of normal and superior children as well, and were improved in many other ways. Many revisions and translations of the Binet scale have been made since then, and various forms of it are now in use in numerous countries.[32]

While it has been generally recognized from the beginning that Binet's tests are not suitable for testing the deaf, two of his basic principles have been accepted in the development of performance tests, as well as language tests, since his time: (1) the use of a group of tests, and (2) the average performance at each age as a standard of measurement. He early recognized the need for testing the "complex processes" to get a true picture of mental ability, and finally rejected specific tests of memory, attention, and other processes, saying that these qualities can not be separated.[33]

The first performance tests were form boards devised by Seguin in connection with the training of mental defectives.[34] Greenberger is credited with being the first to apply psychological testing principles to the deaf, in 1869, although his methods were rough and subjective in the extreme. His tests consisted of: sorting beads and tablets to distinguish form and colors, block building and arranging sticks in patterns.[35]

---

32. Boynton, op. cit., pp. 158-160.
33. Schieffelin and Schwesinger, op. cit., pp. 4-5
34. Drever and Collins, op. cit., p. 18.
35. Greenberger, D., "Doubtful Cases," American Annals of the Deaf, 34: 93-96, 1889

Mott found deaf children to be equal to the hearing on memory and observation tests in 1899,[36] but in a study of 184 deaf children in the Chicago schools in 1906, MacMillan and Brunner found them to be retarded about three years on tests of cancellation, perception of size and weight, memory, and sensory-motor ability.[37]

Up to this time, most of the work on the deaf was theoretical, although the need for an adequate measurement of their intelligence was recognized. In 1912, Kilpatrick suggested the use of the Binet-Simon scale for the deaf, and in 1915, Pintner and Paterson attempted to adapt the Goddard-Binet, but found the average mental retardation of the deaf to be 4.58 years on it.[38] They concluded that it was not a suitable instrument for testing the deaf.

These two investigators next studied the possibilities of using the Digit-Symbol group test, which involved the use of little language, and again found a very large percentage of the deaf rated in the dull class.[39] According to this study, the deaf were retarded about three years, although Newlee's results using the same test, with eighty-five deaf children in the Chicago schools, showed them to be equal

---

36. Mott, A. J., "The Ninth Year of a Deaf Child's Life," American Annals of the Deaf, 44: 401, 1899.

37. Drever and Collins, op. cit., pp. 15-16.

38. Pintner, Rudolf and Paterson, Donald G., "The Binet Scale and the Deaf Child," Journal of Educational Psychology, 6: 201, 1915.

39. Pintner, Rudolf and Paterson, Donald G., "A Class Test with Deaf Children," Journal of Educational Psychology, 6: 591-599, 1915.

to the hearing.[40]

After using several other language tests, Pintner and Paterson concluded that special tests would be required to test the deaf, since their lack of comprehension and environmental experience, and their peculiar psychology handicapped them greatly on the linguistic scales.[41] These authors then worked out a performance scale, using some original tests, and others developed by other investigators. Their criteria in the selection of tests were:

1. To get as many different kinds of tests as possible.
2. Each test should present a relatively new situation to the child.
3. No verbal instructions should be necessary in order to give the tests.[42]

After the standardization of the scale, and the rejection of certain tests which did not give satisfactory norms, Pintner and Paterson published their battery of fifteen tests in 1917. This was revised and shortened to nine tests by Pintner and Spaid in 1918. The results of experiments with this test indicate that the general mental inferiority of the deaf is from two to three years. However, there was no difference between the congenitally deaf, and those whose hearing impairment was acquired.[43] Pintner believes this mental retardation may be due to the same cause that produced

---

40. Newlee, Clara E., "A Report of Learning Tests with Deaf Children," Volta Review, 21: 216, March 1919.
41. Pintner and Paterson, A Scale of Performance Tests, pp. 19-20.
42. Ibid., pp. 21-24.
43. Pintner, op. cit., pp. 318-327.

the deafness.

Within the next few years, several other investigators developed performance tests of various types, but the lack of adequte standardization on most of them has prevented an understanding of the results obtained thus far.

Healy and Fernald devised a group of tests,[44] some of which have been used by later workers, but these were not grouped in a scale nor standardized. Knox constructed a rough scale of performance tests for use with the immigrants at Ellis Island, but it was unstandardized.[45] Some of his items have been used extensively by others, however, and norms have been established on them.

After a survey in Australia in 1918, Porteus concluded that his scale of mazes was suitable for the deaf.[46] Goodenough is responsible for a new type of performance test involving drawing a picture of a man, published in 1926.[47] The chief disadvantage is that the scoring is subjective, and like the Porteus maze, the test measures only one aspect of intelligence.[48]

---

44. Healy, W., and Fernald, G. M., "Tests for Practical Mental Classification," Psychological Monographs, 8, 2, 1911.
45. Knox, H. A., "A Scale, Based on the Work at Ellis Island, for Estimating Mental Defect," Journal of the American Medical Association, 62: 741-747, March 7, 1914.
46. Reamer, Jeannette Chase, "Mental and Educational Measurements of the Deaf," Princeton, New Jersey: Psychological Monographs, 29: 16.
47. Goodenough, F. L., The Measurement of Intelligence by Drawings, New York: The World Book Co., 1926.
48. Boynton, op. cit., pp. 197-199.

Kohs, designer of the Block-Design test,[49] advises its use with other tests, as a check on the Binet. It has been incorporated in several of the recent scales.

In their scales for testing preschool children, Stutsman and Bayley have used a few of the older tests, as well as the new, but some of these items involve the use of language.[50]

Although Ferguson reported high correlation with other estimates of intelligence for his series of six form boards,[51] his findings are of no great significance, since he disregards the factor of age.[52]

Gaw studied intercorrelations of several well-known performance tests, and correlations between them and other standardized intelligence tests, and concluded that for performance tests to be of any value, several different types must be given. She believes that performance tests are superior to the Binet-Simon scale for measuring the intelligence of children with limited environmental experience. To quote her directly:

> The fact that the inter-correlations among the performance tests are rather low makes it very necessary to use not one or even two or three, but a number of tests to arrive at a mental age, just as the Binet scale combines several tests...[53]

Morris, in a recent critical analysis of performance

---

49. Kohs, S. C., Intelligence Measurement, New York: The MacMillan Co., 1923.
50. Boynton, op. cit., p. 202.
51. Ferguson, G. O., "A Series of Form Boards," Journal of Experimental Psychology, 3: 47-58, 1920.
52. Boynton, loc. cit.
53. Ibid., pp. 203-204.

tests, found that they measure three of the seven independent mental abilities identified by Thurstone in his analysis of linguistic tests, namely: Visualizing, Perceptual Speed, and Induction. The traits which were not found to be group factors or common elements in performance tests, but which are included in linguistic tests are Number Facility, Word Fluency, Memory, and Verbal Reasoning. Morris found that the interrelationships among the performance tests selected ranged from a moderately high positive to a moderately high inverse relationship.[54]

In Reamer's study reported in 1921, are the first results based upon a sufficient number of cases of the deaf to be significant. Using the Pintner Non-Language and Educational group tests with twenty-five hundred deaf children, she found that the deaf showed an average mental retardation of about two years, and an average educational retardation of five years.[55]

In 1925 Pintner and Paterson reported that the deaf showed a mental retardation of from two to three years on their non-verbal and substitution tests.[56] These results were supported by the nation-wide survey of schools for the deaf in 1924-1925, made by Day, Fusfeld, and Pintner, under the auspices of the National Research Council. They took a representative sampling of the deaf above twelve years of

54. Morris, Charles M., "A Critical Analysis of Certain Performance Tests," Journal of Genetic Psychology, 54: 85-105, March, 1939.
55. Reamer, op. cit., p. 130.
56. Pintner, Intelligence Testing, pp. 316-327.

age, from thirteen day and twenty-eight residential schools, and gave this group the Pintner Non-Language and Educational Survey tests. Pintner again concluded that the average mental retardation of the deaf is from two to three years.[57]

These studies have been subjected to rather serious criticism by other investigators. McManaway found both low reliability and low validity for the non-language test, and Boynton claims that speed is stressed to too great an extent. He also says that Pintner's selection of schools may have given an undue numerical weighting to the manually-taught deaf, who are generally believed to be inferior, because of the tendency in many schools to select only the brighter pupils for oral training.[58]

Further evidence that the deaf are not as greatly retarded as is generally supposed is produced in the work of Drever and Collins, published in 1928. Their series of performance tests is composed of items from several well-known tests, upon which provisional norms were established. Although their cases number only two hundred, they believe that the deaf are not inferior to the hearing. Their results indicated that in no case were the deaf as much as one year retarded, and at the eleven and twelve year levels, the deaf were superior to the hearing.[59]

Undoubtedly the chief cause of the conflicting results

57. Pintner, Rudolf, "The Survey of Schools for the Deaf, V. Psychological Survey," American Annals of the Deaf, 72: 273-299, 1927.
58. Boynton, op. cit., pp. 438-440.
59. Drever and Collins, op. cit., pp. 17-52.

on the relative intelligence of the deaf, when the same test is used, is the fact that the scales have been standardized on totally different populations. In 1930 Arthur published her point scale of performance tests, standardized on eleven hundred public school children from middle-class American homes. While all of the items are from other well-known tests, the author stipulates that the order of presenting them must be rigidly observed, since this factor will affect the norms.[60] Her chief difficulty in constructing the scale was to find items sufficiently difficult to test the superior adult, without involving school training or experience, and she recommends that further studies be made on this problem.[61]

In 1933 MacKane made a comprehensive investigation to determine whether or not the findings of the Drever and Collins study were correct, since they differed so greatly from those of the Hesmer and National Research Council studies. He matched one hundred thirty deaf subjects between the ages of ten and twelve with hearing children, for chronological age, sex, socio-economic status, and nationality of parents. He used the Drever-Collins, Pintner-Paterson, and Arthur scales, but since some of the subtests appear on all three scales, they could not be given separately. He combined the items from all three in a battery, preserving the order of the Arthur test so that the norms would be affected as

60. Arthur, <u>op. cit.</u>, pp. 1-82.
61. Arthur, Grace, <u>A Point Scale of Performance Tests</u>, New York: The Commonwealth Fund, Division of Publications 1933, 2: 201.

little as possible. No speech was used by the examiner in giving the test to the hearing or the deaf. Finally, the Pintner Non-Language test was given to the same children, to serve as a check. MacKane's results on the battery of tests supported those of Drever and Collins in general, since he found that at no age level were the deaf as much as one year retarded. However, he found no superiority of the deaf. The Pintner Non Language test showed a marked superiority of the hearing, which convinced MacKane that it does not measure the same abilities as the performance scale.[62]

In 1936 Peterson reported the results of testing one hundred deaf children on the Kohs Block Design, and concluded that it is the most satisfactory single performance test yet designed.[63] He found the degree of retardation of the deaf to be less than on the Goodenough test, which could be expected, since the positive correlation between the Block Design and the Binet is .64,[64] while that of the Goodenough and the Binet is .76.[65]

After Eigler reported finding several abnormally high Intelligence Quotients on the Kohs test with preschool children, Benton gave it to a similar group, but found no great

62. MacKane, Keith, "A Comparison of the Intelligence of Deaf and Hearing Children," New York: Teachers College, Columbia University, 1933, Contributions to Education, 588: 6-44.
63. Peterson, Edwin G., "Testing Deaf Children with Koh's Block Designs," American Annals of the Deaf, 81: 242-254, May, 1936.
64. Kohs, op. cit., p. 157.
65. Goodenough, op. cit., Introduction.

discrepancy between it and the Binet.[66]

Springer reports on a survey in 1938, in which three hundred thirty matched pairs of hearing and deaf children were given the Goodenough test. Although the hearing children were slightly superior, the mean Intelligence Quotient of the deaf was 96.24, indicating very little retardation.[67]

In a recent survey made by Bowers at the West Virginia School for the deaf, the mean Intelligence Quotient of the two hundred pupils tested was 89.2. Of the nine performance items given the lowest mean rating, 69.29, was on the Mare and Foal test, while the highest, 101.25, was on the Porteus Maze.[68]

In the individual testing program at Central Institute for the Deaf, Lane reports a median Intelligence Quotient of 97.6 for the two hundred fifty cases given the Randall's Island series, a highly standardized and reliable test for the preschool age child.[69, 70] The Lectometer, designed by Meyer,[71] gave results correlating highly with the Binet on

66. Benton, Arthur L., "The Performance of Pre-School Children on the Kohs Block Design Test," Journal of Genetic Psychology, 53: 231-[illegible], Sept., 1938

67. Springer, N. Norton, "A Comparative Study of the Intelligence of a Group of Deaf and Hearing Children," American Annals of the Deaf, 83: 151, March, 1938.

68. Roth, Stanley D., "Survey of the Psychological Examination Given by Dr. Stella M. Bowers, May 1937," The West Virginia Tablet, 61, 1-6, April, 1938.

69. Lane, op. cit., pp. 170-171.

70. Schick, Helen F., "The Use of a Standardized Performance Test for Preschool Age Children with a Language Handicap," Proceedings of the International Congress on the Education of the Deaf, June, 1933, pp. 526-532.

71. Schick, Helen F., and Meyer, Max F., "The Use of the Lectometer in the Testing of the Hearing and the Deaf," American Annals of the Deaf, 77, pp. 292-304, Sept., 1932.

hearing children. This instrument was used for testing two hundred older children, who showed ability equal to that of the hearing. Since the distribution of Intelligence Quotients follows a normal curve, with a range from feeble-minded to genius, and the median quotient falls around 100, Lane concludes that the deaf are not retarded. She says that the retardation reported may be due to the following factors:

1. The use of tests which are not strictly non-verbal.
2. Testing in large groups.
3. Testing by examiners not familiar with the deaf child.[72]

Although the recent studies show less retardation for the deaf than the pioneer work indicated, the difference in results is still too great to justify the drawing of any significant conclusions. Only after the existing tests have been thoroughly standardized and compared to the most reliable of the language tests, can it be definitely decided whether or not the deaf as a group are mentally equal to the hearing.

---

72. Lane, op. cit., pp. 170-171.

CHAPTER III.

PROCEDURE

Description of the Tests

All of the items used in the scale have been previously standardized in the form in which they are given here. The material used is obtainable from the manufacturers of testing supplies. The tests selected are all included in the well-known performance scales. The sequence of presenting the items is that of Drever and Collins,[73] with the Seguin board added in the order of the Arthur scale.[74] The method of scoring is in each case the same as that used by the originator of the test. Since the absence of verbal directions was one of the criteria in the selection of the tests, all of the directions may be given in pantomime. Nevertheless, with the hearing child, they are given verbally, for the sake of naturalness, but are identical in all other respects. The time required to give the test varies from thirty-five to seventy minutes, with an average of about forty-five minutes. A description of the test material, and method of presenting and scoring the tests follows:

Test 1 - Block Design

This material, designed by Kohs[75] to test mental analysis and synthesis, is included in the Drever and Collins,

---

73. Drever and Collins, op. cit., pp. 31-41
74. Arthur, op. cit., 1: 13-30.
75. Kohs, loc. cit.

Arthur, and MacKane scales. It consists of sixteen one inch wooden cubes, painted in the same manner, the sides being white, blue, red, yellow, red and white divided diagonally, and blue and yellow divided diagonally. Seventeen designs, printed in the same colors on small cards, are presented in increasing order of difficulty for the subject to construct.

The procedure of Drever and Collins is followed to explain the test and to insure a thorough understanding of the task. Four of the sixteen blocks are placed before the subject. The examiner turns each of the blocks until all four have the blue side up. Then the blocks are turned to show successively the remaining faces: red, yellow, white, red and white, and blue and yellow. The examiner turns the first block and the child matches the remaining three to this. The object of this procedure is two-fold, namely: to show the child all the faces of the blocks, and to make clear to him that all blocks are the same. The examiner then constructs the first design as a sample, scatters the blocks, and indicates that the child should put them together. If he succeeds in accomplishing the completion of the design within the assigned time limit, he is permitted to go on to the next design. The scoring depends upon two variables, the number of moves, and the time taken for each design.

Test 2-Knox Cube

This test for memory span was devised by Knox,[76] and is used, either in the original form or with slight variations, by Drever and Collins, Pintner and Paterson, Arthur, and MacKane. The material consists of six one inch wooden cubes, painted black. Four of these are placed before the subject in a row, about two inches apart. The remaining two are used by the examiner and the subject for tapping. Some investigators use only one cube for tapping, passing it back and forth from examiner to subject. The procedure necessitates the loss of time and distraction of the pupil's attention, occasionally causing his failure on a test which he might have passed otherwise. Another slight variation from the standard procedure is the introduction of a sample pattern, 1-2-3, to be sure the subject understands that he must imitate the pattern tapped by the examiner. This sample pattern, used by Maxfield of Ohio State University, is a substitute for the verbal directions given in testing manuals.

"Watch carefully, and then do as I do. Do that."
The examiner taps a pattern with his cube, at the rate of one tap per second, and indicates that the subject is to imitate the pattern. The twelve patterns used in this scale are arranged in the order used by Drever and Collins, which is slightly different than that used by Pintner. The subject is allowed three failures before the test is discontinued. The

76. Knox, <u>loc. cit.</u>

test is scored according to the number of patterns successfully completed.

Test 3-Seguin Board

The Seguin Board test, standardized by Sylvester, was used by Pintner and Peterson, Drever and Collins, Arthur and MacKane.[77] It consists of ten blocks of common geometrical shapes, which are to be placed in appropriate places on the board at a given signal. Three trials are given, and the shortest time is recorded. The arrangement of the blocks in three piles at the beginning of the test, in the manner specified by Sylvester, must be adhered to if the norms are not to be affected.

Test 4-Manikin and Profile

Both of these tests for synthesis were included in all four of the well-known scales. The Manikin test, devised by Pintner,[78] is used for younger children, as he has suggested. The material consists of a conventionalized figure of a man, cut from wood, and divided into six pieces -- the head, the body, the two arms, and the two legs. The examiner arranges the pieces in the manner described by Pintner and indicates that the subject is to put them together, without telling him what the finished product will represent. A maximum time limit of five minutes is allowed, and the test is

77. Pintner and Peterson, A Scale of Performance Tests, pp. 30-33.
78. Ibid., pp. 53-55.

scored according to the accuracy of the assembled figure. Perfect performance scores a mental age of seven years, and therefore is not applicable to older children. If the child scores a perfect performance on this test, he is given the Feature Profile, and the earlier test is simply recorded as completed.

The Feature Profile test, devised by Knox,[79] consists of a human head cut out of wood, and divided into eight pieces. The time limit is five minutes, and the scoring is based on the time required to complete the test. The subject is not told what the figure is intended to represent, in accordance with Pintner and Paterson's procedure.

Test 5-Form Boards

This test is composed of a series of three form boards: the Two-Figure, the Healy Puzzle A, and the Casuist. Each test is scored individually, and the average of the three is taken for the score on the test. Five minutes is allowed for each board, and the score is determined on a basis of time required for completion.

The Two-Figure board,[80] devised by Pintner, is used in all of the standardized scales. The board has two cut-outs, a square and a cross, into which must be fitted the nine pieces arranged above the board.

The Healy Puzzle A, designed by Healy and included by

79. Knox, op. cit., p. 741.
80. Pintner and Paterson, op. cit., pp. 35-37.

Drever and Collins, MacKane, and by Pintner and Paterson in the longer version of their performance series, is of the specifications given by Pintner and Paterson.[81] It consists of a frame, into which five rectangular pieces are to be fitted.

The Casuist Board, used in all of the scales except that of Drever and Collins, was devised by Knox.[82] It consists of a rectangular board with four spaces, three circles of varying sizes and an oval, into which twelve pieces must be fitted.

Test 6-Picture Completion

Healy's Picture Number 1[83] is used in each of the four standardized scales. It is an attractive, colored picture of an outdoor scene, in which ten different activities are represented. In connection with each of these, a one inch square is cut out. The subject is to select from fifty squares, the one most suitable for insertion. Although the test is usually finished in five minutes, ten may be allowed if necessary. The score is determined only by the correctness of the solution. There are several possible solutions for each insert, but only one correct one which gives full credit. However, the selection of any of the other possible choices allows a few points credit on the total score. Mental age is determined from the total

81. Ibid., pp. 44-53.
82. Knox, op. cit. pp. 741-747.
83. Pintner and Paterson, op. cit., pp. 61-63.

number of points, following Healy's norms.

Test 7-Drawing Test

This test, adapted from the Binet Scale,[84] is a supplementary test used as a check with subjects whose ability has not exceeded a seven year average on the preceding tests. The examiner shows a card on which a circle has been drawn, and indicates that the subject is to copy three such patterns from the model. If he succeeds in drawing one acceptable circle in the three trials, he continues in the same manner with the square, triangle, inverted triangle, and diamond, until he fails one of the tests. The test is scored in the following manner:

Circle.....3 years
Square.....4 years
Triangle.....5 years
Inverted triangle....6 years
Diamond.....7 years

When the mental age on each of the six (or, with the drawing test, seven) items of the scale has been computed, the mental age is determined by finding the mean of these scores.

The validity of the scale was determined by giving the test to 209 children who had been given other standardized and reliable tests. The results were compared with the scores on these tests, and correlation coefficients were computed.

84. Terman, Louis M., and Merrill, Maud A., Measuring Intelligence, Boston: Houghton Mifflin Co., 1937, p. 201 219,231.

There were 133 deaf and speech defective cases at Central Institute for the Deaf who had been given other performance tests. Sixty-five were tested on the Randall's Island scale for younger children, and sixty-eight on the Lectometer designed by Meyer.

In the comparison of the performance scale with language tests, the scores used were those obtained by examiners in the public schools attended by the 106 hearing children. The linguistic tests used for comparison were: Stanford-Binet, Detroit Kindergarten and First Grade, Haggerty, Henmon-Nelson, Kuhlmann-Anderson, Otis, and Pintner Cunningham.

The hearing group tested includes pupils enrolled at the Central Institute Speech Correction Clinic, who are attending public and parochial schools in St. Louis and its suburbs, and nearby towns in Missouri and Illinois. In order to secure greater numbers of cases, the author obtained permission to do additional testing at the Arlington School in St. Louis, the Lockwood School in Webster Groves, and the Temple Israel Sunday School.

The author also used the scores of the 106 hearing cases and the 133 deaf and speech defective cases in determining the variation between the scores on the performance scale and those on language and other performance tests.

In addition to the pupils whose test scores were used in the comparison of the tests, other speech defective and deaf subjects have been tested with the performance scale.

These scores were included in the comparison of the performance of the hearing, deaf, and speech defectives.

Results of this study are presented in the form of tables. The statistical procedure involved the following steps:

The correlation coefficients between the performance scale and other performance tests, and between the performance scale and linguistic tests, were computed by means of the gross score formula:

$$r = \frac{N\Sigma X_1X_2 - \Sigma X_1 . \Sigma X_2}{\sqrt{N\Sigma X_1^2 - (\Sigma X_1)^2}\sqrt{N\Sigma X_2^2 - (\Sigma X_2)^2}}$$

The probable errors were found by using the following forumula:

$$P.E._r = \frac{.6745\ (1-r^2)}{\sqrt{N}}$$

The variation between the performance scale score and the language test score for each case was tabulated according to the number of points difference, and the percentage of the total number of cases showing this variation was recorded for both language and performance tests.

The mean and median Intelligence Quotient were found for the deaf, hearing, and speech defective children, as measured by this performance series.

The mean and median Performance Quotients of the deaf

were found for each test item included in the scale, to determine which tests were most difficult for the group, and which tests were easily completed. In finding the scores of the deaf on separate performance items, the scores of all deaf children who have been tested on this scale so far, were included.

CHAPTER IV.

RESULTS

Feeling the need for a performance series that would measure the intelligence of deaf children of school age, Lane assembled performance tests from various sources into such a series. Mental ages were determined from the norms established for hearing children on the separate tests, and the average of the groupings was used as the mental age on the series. Although Lane has tested deaf children and speech defective children using this series, no study of the validity of the test had been made. The results of this investigation will attempt to show the relationship between scores on this performance scale and other standardized tests of intelligence.

Table 1 gives the results of the comparison made between the scores of the deaf and speech defective children on the performance scale and on the other performance tests. The probable errors are very low, indicating high statistical reliability for the test. The correlation coefficients, which are fairly high, show that this scale measures the same type of intelligence measured by the other standardized performance tests.

It might be expected that the correlation would be higher with the Randall's Island series than with the Lectometer test, since the former is more nearly the same type of test as the performance scale. But it must be remembered that the Randall's Island test is given the

preschool children at Central Institute, often upon their arrival at the school, when the strangeness of a new envir-

Table 1.

Comparison of the Performance Scale with Other Performance Tests

Test	N	r	P.E.$_r$
Lectometer	68	.78	± .03
Randall's Island	65	.65	± .04
Total Performance Tests	135	.71	± .03

onment, the fatigue due to a long trip, and the many behavior maladjustments typical of the young deaf child before he attends school, all affect his performance on the test. In cases where a retest was made the following year, the score was often altered considerably, not because of any shortcoming in the test, but rather because of the improved habits of attention and concentration of the subject, as well as of his improved behavior. In all cases on which a retest was made, using the same test, the second score was used in the comparison with this performance series. However, in many cases, the child was ready for the Lectometer or advanced performance series by this time, so the original score was the one used in the correlation.

Another reason for the greater discrepancy in scores between the Randall's Island and the performance scale might be the greater number of years intervening between the giving of these two. Since the Randall's Island was usually

followed by the Lectometer test, several years frequently elapsed between the giving of the former test and the performance series. Although the Intelligence Quotient should remain constant throughout an individual's life-span, if properly measured, many of the scores of language tests obtained on the hearing children also showed an increase in the Intelligence Quotient on subsequent tests, probably due largely to the unfamiliarity of the subject with a testing situation on the first test.

In Table 2, the comparison is shown between the scores of the hearing children on the performance scale, and on the language tests. The probable errors of the correlation between the total language tests, and each separate test except the Kuhlmann-Anderson, with the performance series, show statistical reliability. While the validity coefficients in general are not as high as they are on the other performance tests, they still show a fair degree of relationship between the two types of tests. The language tests with which this scale is compared show an average validity coefficient of .715 with the Binet, with the individual correlations ranging from .55 to .85. The average validity coefficient of language tests with the Otis test is .735, and the range is from .66 to .81.

Table 2.

Comparison of the Performance Scale with Language Tests			
Test	N	r	P.E.$_r$
Binet	26	.65	± .08
Henmon-Nelson	21	.66	± .08
Kuhlmann-Anderson	21	.19	± .14
Detroit	15	.56	± .12
Total Language Test	106	.63	± .04

Only the four tests on which the greatest number of scores were obtainable were used for individual comparisons, since the small number of cases on the Haggerty, Otis, and Pintner-Cunningham did not justify statistical treatment. However, these cases were included in computing the correlation of the total language tests.

The Detroit Kindergarten and First Grade Tests might be expected to show a lower degree of correlation than the other language tests for the same reason which was advanced for the Randall's Island series. The date of giving the other language tests more nearly approximates that of giving the performance test, than does the date of the Detroit test.

The Binet scores might have shown a somewhat different relationship, were it not for the selection of the cases, although this is only a conjecture on the part of the author. Eleven of the cases at the Temple Israel Sunday School were given the Binet and performance tests because

they were considered problem cases, and their teachers felt that psychological diagnosis was needed. All of the subjects were either behavior problems or poor readers. Most of the other Binet scores were obtained from the public schools, which give an individual test only when the pupil is regarded as a problem, either because of behavior or poor scholarship. Therefore, the author feels that it is safe to conclude that the children tested on the Binet were not an average group.

No explanation can be offered for the great difference in the results of the Kuhlmann-Anderson test, other than the fact that the tests may not be measures of the same kind of intelligence, or that the Kuhlmann-Anderson was not given following standard procedure. Another point worth consideration is that this language test was not standardized by comparing it with other tests, since the authors do not endorse this type of standardization.

On a whole, the comparison of the performance scale with other performance tests and with language tests, shows a high degree of reliability. Although the correlation coefficients are not sufficiently high to justify making any claims as to the prognostic value of the scale, they do indicate that the tests are measuring the same type of ability to a fairly great extent. Within certain limits, they could be used to good advantage for purposes of homogeneous grouping, and a fairly accurate estimate of intelligence. Their chief use with the normal hearing

individual would be as a check on the linguistic test. Their value in testing the deaf and speech defective would undoubtedly be increased by being given in conjunction with another performance test of a different type.

Table 3 shows the variation, in terms of points, between the scores of the performance scale, and the Intelligence Quotients on language tests and other performance tests. In such a comparison, it must be borne in mind that the scores on two successive givings of the same test fluctuate to a certain extent. Lane reports that on language tests

> .....The fluctuation of from four to seven points is considered normal, due to environmental conditions, health of the child, attention, concentration, and to chance factors.[85]

However, in a recent paper read by Stoddard at a meeting of the Educational Records Bureau, he reported a rise of twenty points in Intelligence Quotients of some nursery school children. He also reported a change of average Intelligence Quotient of from 87 to 116, in a group of illegitimate children after placement in foster homes. There was a period of from one and one-half to six years between tests.

---

85. Lane, Helen Schick, "The Stability of Mental Test Ratings of Preschool Age Deaf Children," Oralism and Auralism, p. 30, Feb., 1938.

Table 3.

Comparison of Scores on the Performance Scale with Scores on Language and Other Performance Tests		
Variation	Language Tests	Other Performance Tests
Less than 5 points	23.6 per cent	27.1 per cent
Less than 10 points	47.2 per cent	57.1 per cent
From 10 to 19 points	34.0 per cent	29.3 per cent
Over 20 points	18.8 per cent	13.5 per cent

In view of this evidence on the inconstancy of the Intelligence Quotient, the fact that the comparisons of the performance scale scores with almost half of the language tests, and over half of the other performance tests, came within a range of ten points is surprising. This fact takes on added significance when one realizes that in many cases where two or more language scores were reported on the same pupil, the range between them was considerably greater than ten points. This study furnishes additional evidence that the performance scale tests the same type of ability as the other standardized tests.

Table 4

Comparison of Hearing, Deaf, and Speech Defectives on the Performance Scale					
	N	Range of C.A.	Mean I.Q.	Median I.Q.	Range of I.Q.
Normal Hearing	47	7-0--13-7	112.75	113.33	81-163
Deaf	115	5-7--21-4	103.57	104.07	60-141
Hearing with Speech Defects	105	6-0--19-10	97.12	99.00	40-140

Table 4 makes a comparison between the mean and median Intelligence Quotients of the hearing, deaf, and speech defective. The range of scores of the speech defectives is somewhat wider than the ranges of the other two groups, because of a few extreme cases in the Speech Correction Clinic, whose speech defect is caused by feeblemindedness. It should be pointed out that the mean is also influenced by the inclusion of these cases, as well as of a few spastics, in both the deaf and speech defective groups. The spastic child, as well as several deaf children whose hearing impairment involves the loss of vestibular stimulation, show poor muscular coordination, and the score on a performance test is not an accurate measure of their mental ability.

The author believes that the majority of the hearing cases used in this comparison were a select group, since the choice of pupils to be tested in the public schools was necessarily left to the principals and teachers. Almost invariably, the outstanding pupils in the class were selected for the test. On the other hand, the cases of behavior maladjustment given the Binet might tend to balance this selection, although many of them scored higher on the performance test than on the linguistic test, because of retarded language ability.

The highest score among the hearing requires some explanation, since the child who scored 163 on the performance scale had an Intelligence Quotient of 96 on the Binet test. This eight-year old boy was tested at the request of

his Sunday School teacher, who reported that he appeared to be very intelligent, in spite of his great difficulty with reading, and his unusually small stature. Although he was rather shy, this boy displayed great interest in the puzzle-solving items of the test, and seemed to have a definite ability along this line. Nevertheless, he demonstrated very marked reversal tendencies on the Koh's Block Design test, which would account for his reading difficulties, and perhaps in part for his poor performance on a language test.

The deaf group represents the most normal distribution of the three classes, since it is composed of all the deaf pupils at Central Institute who have reached the age level of the test. This is an average group of deaf children, presumably, although some educators of the deaf might raise the objection that the deaf capable of receiving oral instruction are of superior ability. If the study were being made in a state school, where mentality is often the basis of determining the method of instruction, the argument might be a valid one. But in a school where only the oral method is taught, all types of pupils are attracted, although there is some selection in the enrollment.

The lower extreme of the range for the deaf represents the score of a totally deaf boy who is losing his vision, and has poor muscular coordination as the result of rickets and other serious childhood diseases. On the other hand, the highest score is that of a twelve-year old girl who is doing fifth grade work at school. Between these two extremes lie

cases of all degrees of intelligence, such as one would find in a school of hearing children.

The speech defectives, too, represent all levels of intelligence, although several cases of spasticity and feeble-mindedness affect the group mean.

In spite of the fact that the data derived from this study indicated a superiority of the normal hearing over the deaf, the insufficient number of cases and the selection of the hearing subjects makes the drawing of any significant conclusions unsafe at the present time. Only after the test has been given to far greater numbers of unselected hearing and deaf children, can a comparison be made of the intelligence of the two groups.

Figures 1 and 2 show graphic comparisons of the distribution of Intelligence Quotients of both deaf and hearing. The scores of the speech defective children, who all have normal hearing, were added to those of children tested in the public schools to form the hearing group. In this way, the author hopes to find a more nearly normal distribution of the hearing, since the two groups are slightly weighted at opposite extremes of the scale.

These curves for deaf and hearing are very similar, and indicate normalcy for both groups. The range and standard deviation of the hearing group shows greater scatter, since some of the feebleminded speech cases are included, as well as the selected public school pupils. With the deaf, the mean and median are very nearly the same, but there is a difference

of 2.37 points between these two measures of central tendency in the case of the hearing. The fact that the median is greater than the mean in both cases indicates that a few extremely low scores affected the latter measure in each instance. The negative skewness of both curves also shows a tendency for the scores to be massed at the high end of the scale. This may be explained by one of two causes, or perhaps by both: (1) technical faults in the construction of the test, making it too easy for the hearing group especially, or (2) the special selection of the group tested.

The standard deviation of the distribution for the deaf is 16.65, while that for the hearing is 21.84, which indicates greater variability of the hearing scores.

Figures 1, 2, and 3 are plotted with intelligence quotients in step intervals of ten along the abscissae, and frequency of occurrence of these quotients on the ordinates. Figure 3 is a comparative line graph which shows the distributions plotted in figures 1 and 2, superimposed for comparison.

Figure 1

Distribution of the Intelligence Quotients of the Deaf

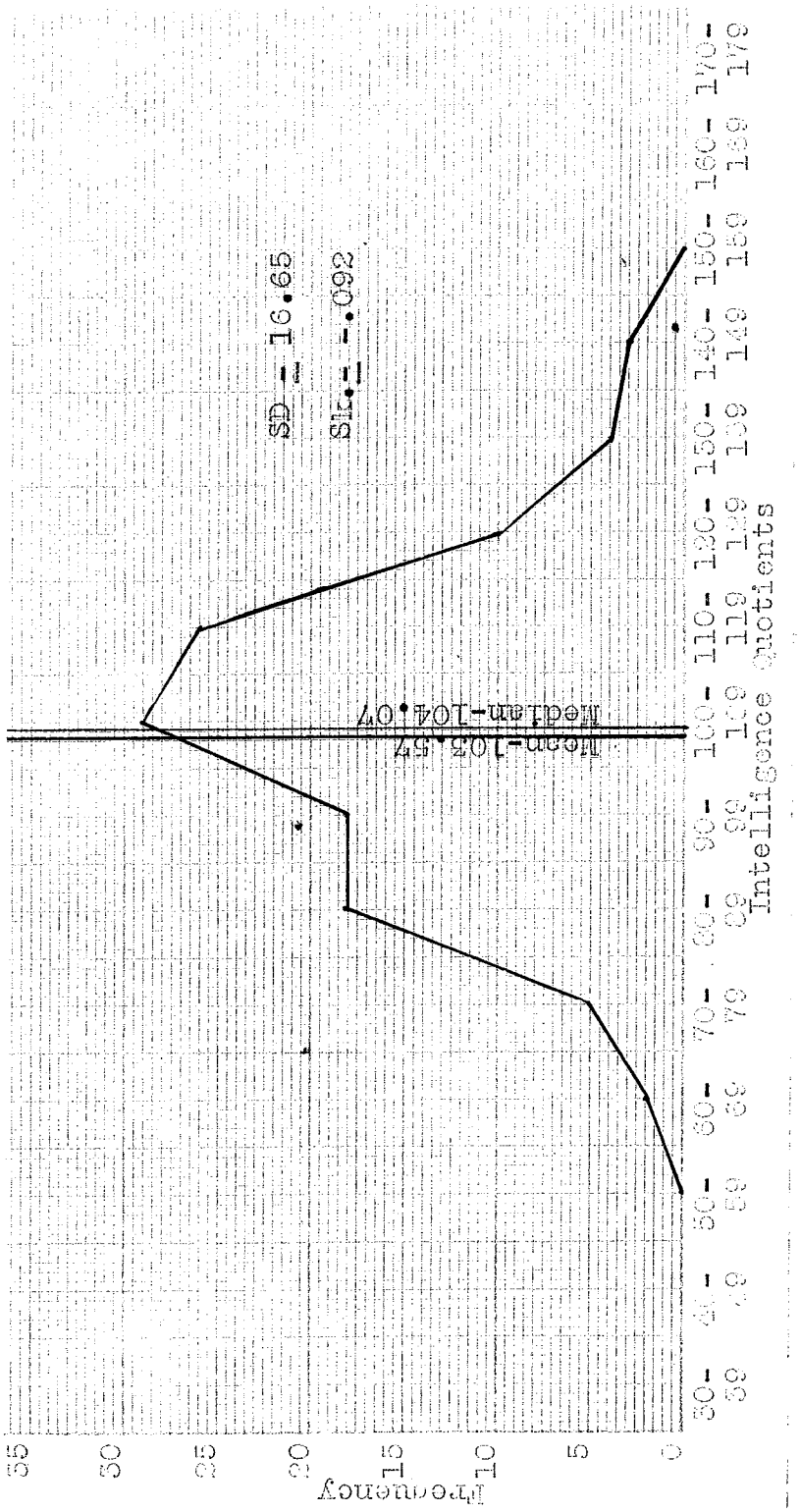

Figure 2

Distribution of the Intelligence Quotients

of the Hearing

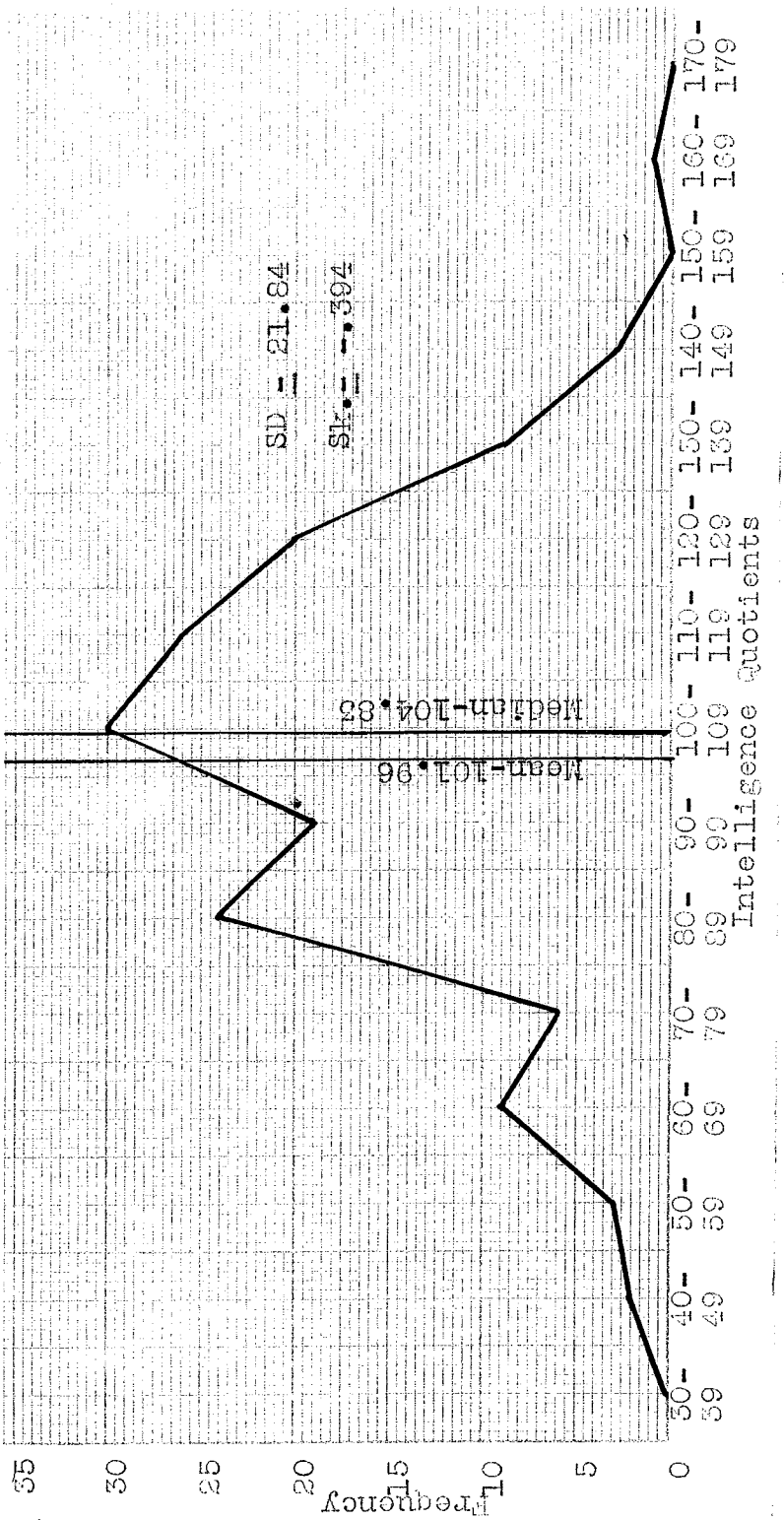

Figure 3

Comparison of the Intelligence Quotients of Deaf and Hearing

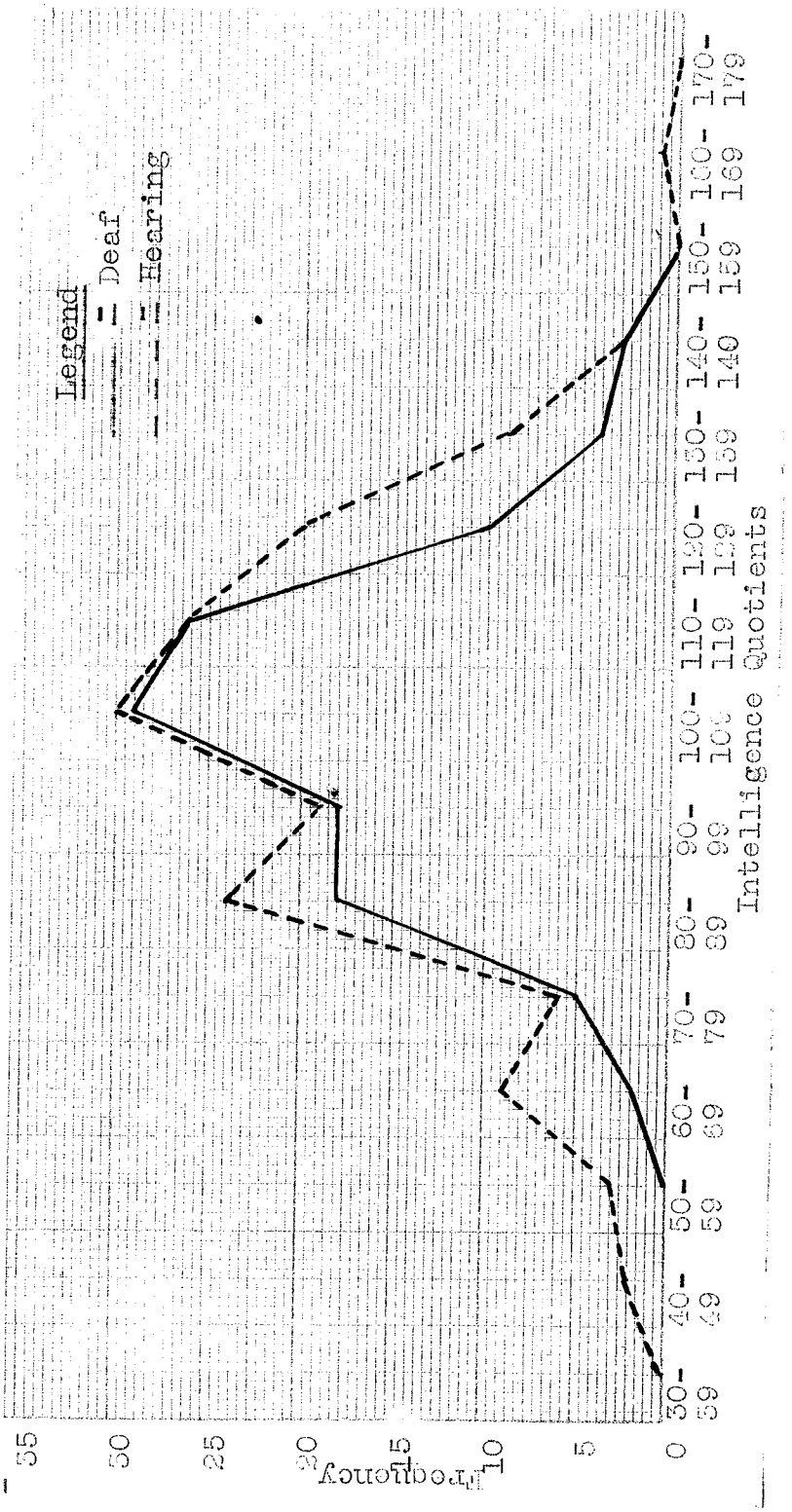

Figure 5 shows the frequency distribution of Intelligence Quotients for both deaf and hearing. From this figure the similarity of the two groups can be noted, as well as the greater variability of the hearing group.

Table 6 gives the mean and median performance quotients of the deaf on each item of the test. The fact that the three tests in which the deaf excel, namely, the Block Design, Manikin-Profile, and Picture Completion, are tests of visual perception, analysis, and synthesis, would seem to indicate that the deaf are above average in these traits. This might be influenced partially by the training of these powers in lipreading and speech.

While the Form Boards also measure visual perception, the fact that they are of a more abstract nature, lacking the colors of the Block Design and the recognizable shapes of the Manikin and Profile, may explain the poor performance of the deaf, whose early training is chiefly concrete in nature. Since the Seguin Board is primarily a test of speed in hand-eye coordination, anyone dealing with deaf children can readily understand the reason for their deficiency in this quality. Their entire education has stressed perfection of execution of every task, at the expense of speed, and it is very difficult to impress upon the deaf the importance of speed on the test. While this acquired trait of the deaf undoubtedly hampers their performance on all of the items, speed plays a much greater part on the Seguin Board than on any of the others, so it is to be expected that they should

do the poorest on it. The inclusion of a few spastic cases, as well as several deaf whose muscular coordination is retarded because of the loss of function of the static labyrinth, would also influence the score on the Seguin Board and other form boards to a great extent.

Table 5.

Scores of the Deaf on Separate Performance Items

N = 115

Test	Range	Mean Performance Quotient	Median Performance Quotient
Block Design	53-173	106.03	102.68
Knox Cube	38-193	95.55	92.87
Seguin	43-148	93.52	88.71
Manikin-Profile	65-206	118.79	116.94
Form Board	47-159	96.56	93.00
Picture Completion	59-253	114.62	113.66
Drawing	33-125	90.72	96.50
Total	60-141	103.57	104.07

The data on the Drawing test is based on too few cases to be very significant, since only eighteen of the deaf children were given this test. However, it seems to give additional evidence that the deaf are retarded in the development of visual perception. The low score on the Knox Cube test indicates a deficiency in memory span, if this test can be considered a reliable index of that ability.

CHAPTER V.

CONCLUSIONS

1. Although several investigators have contributed a great deal of information on the intelligence of the deaf through their studies, the question of whether or not the deaf are mentally retarded is still unsolved. It is necessary, therefore, for other workers with the deaf to collect data, which may eventually lead to a solution of this problem confronting educators of the deaf.

2. For this purpose, standardized and reliable tests must be used. While several are available, few are practical for testing the deaf child of school age.

3. The Performance Scale presented in the present study was compiled by Lane of Central Institute for the Deaf, where it has been in use for several years. It consists of seven tests: Koh's Block Design, the Knox Cube, the Seguin Board, the Manikin and Profile, the Form Boards (Two-Figure, Healy A, and Casuist), the Healy Picture Completion No. 1, and the Drawing Test. Although each of the items has been standardized previously, the scale as a whole has not been standardized.

4. The present study makes a comparison between the scores of 106 public school hearing children on this scale, and on standardized language tests, and between the scores of 133 deaf and speech defective children at Central Institute on this scale, and on standardized performance tests. The language tests used in the comparison were: Stanford-Binet, Detroit Kindergarten and First Grade, Haggerty, Henmon-Nelson,

Kuhlmann-Anderson, Otis, and Pintner Cunningham. The other performance tests compared with the scale are: Randall's Island Performance Series, and the Lectometer.

5. The validity coefficient between the performance scale and other performance tests is .71 ± .03, and that with the language tests is .63 ± .04, both of which are statistically reliable. While the correlations are not sufficiently high to indicate that the test has prognostic value, they may be interpreted to mean that the performance scale measures the same type of ability as the other tests measure, to a large extent.

6. The fact that 47.2 per cent of the language test scores, and 57.1 per cent of the other performance test scores, vary less than 10 points from the scores on the performance scale, appears to give additional evidence that the latter is a valid instrument for the measurement of intelligence.

7. The comparison of the three groups tested on the scale shows the hearing to be decidedly above normal, with a median Intelligence Quotient of 113.33. The author believes that the selection of the cases may have some influence on this score. The median of the deaf, presumably an average group, is 104.07. The fact that the entire group is orally-taught may indicate that it is slightly above average. The speech defective group has a median of 99.00, but the inclusion of the scores of several spastic and feebleminded children who have speech defects may have altered this group's rating. From this data, it appears that all three groups are within

the normal classification. The insufficient number of cases and lack of matching make a statement regarding the relative intelligence of the hearing, deaf, and speech defective groups unjustified at this time.

8. The scores on separate performance items revealed that the deaf rank above average on the Block Design, Manikin-Profile, and Picture Completion tests. Since these items are intended to test synthesis, analysis, and visual perception, it is apparent that the deaf are not deficient in these traits.

9. The deaf were found to be below normal on the Knox Cube test, which may be interpreted to mean that they are retarded in memory span.

10. The low scores on the Seguin Board seem to indicate that the deaf are below average in perceptual speed, probably because speed is not stressed in their education. These low scores may also be due to poor motor coordination caused by spasticity, or the loss of the vestibular function.

11. The Form Board scores show a retardation in visual perception of an abstract nature. This deficiency may be attributed to training also, or to poor muscular coordination.

12. The results as a whole reveal that the deaf are probably normal in intelligence, and certainly not as greatly retarded as the studies of several investigators would indicate.

13. In evaluating the scale from the standpoint of its future use, a few points merit consideration. The excessively high scores on the Manikin-Profile and Picture Completion

suggest that a verification of the norms of these tests might not be out of order.

14. The test appears to give undue weighting to the visual perception aspect of intelligence, but until further studies of the nature of intelligence have been made, and the qualities tested by the various items have been more clearly defined by factorial analysis, the discussion of this question is unjustified.

The author concludes that the scale as a whole may be recommended as a reliable and valid instrument for testing the deaf and speech defective child of school age. Its value with both groups would be greatly enhanced by confirming the findings on another reliable performance test of a different type. Its use may also be recommended with the normal hearing child as a check on the linguistic test. The chief advantages of the scale are: (1) the variety of intelligence traits measured, (2) the complete absence of oral instructions and responses, (3) the comperative ease of administration and scoring, (4) the time required for testing, (5) the attractiveness of the tests, and (6) the portability of the material.

The evidence from this study warrants the recommendation that a testing program be installed in every school for the deaf. While the first test may give an accurate estimate of the child's intelligence, many cases on both language and performance tests indicate that it does not always do so. Moreover, this testing program should include a variety of tests, since the variation in scores on different tests shows

a difference in the traits tested by the various measures.

One disadvantage which many school administrators mention is the cost of the testing material. While the initial cost is fairly high, the material is practically indestructible, and the average cost per pupil over a several year period would not exceed that of the "paper and pencil" tests which are sold by the publishers of testing materials. The material necessary for this test is also considerably less expensive than that required for most of the well-known performance scales. Since one examiner would probably do all testing in a school, the investment in one set of equipment would not be excessive.

It is hoped that the beginnings made in this investigation will lead to the collection of further data, by the use of which more accurate information on the nature of intelligence and its measurement with the deaf may be obtained. The great need for this work in the education of the deaf is widely recognized, and future investigators are urged to apply their efforts "to help the handicapped child to help himself."[86]

---

86. Goldstein, M. A., Lecture notes, 1938.

BIBLIOGRAPHY

Arthur, Grace, A Point Scale of Performance Tests, I, Clinical Manual, New York: The Commonwealth Fund, Division of Publications, 1930.

_____, A Point Scale of Performance Tests, II, The Process of Standardization, New York: The Commonwealth Fund, Division of Publications, 1933.

Benton, Arthur L., "The Performance of Pre-School Children on the Koh's Block Design Test," Journal of Genetic Psychology, 53: 231-235, Sept., 1938.

Boynton, Paul L., Intelligence, New York: Appleton and Co., 1933, pp. 1-30, 145-204, 426-441.

Bronner, Augusta F., Healy, William, Lowe, Gladys M., Shimberg, Myra E., A Manual of Individual Mental Tests and Testing, Boston: Little, Brown and Co., 1932.

Drever, James, and Collins, Mary, Performance Tests of Intelligence, Edinburgh: Oliver and Boyd, 1928.

Ferguson, G. O., "A series of Form Boards," Journal of Experimental Psychology, 3: pp. 47-58, 1920.

Goodenough, F. L., The Measurement of Intelligence by Drawings, New York: The World Book Co., 1926.

Greenberger, D., "Doubtful Cases," American Annals of the Deaf, 34: 93-96, 1889.

Healy, W., and Fernald, G. M., "Tests for Practical Mental Classification," Psychological Monographs, 8, 2, 1911.

Knox, H. A., "A Scale, Based on the Work at Ellis Island, for Estimating Mental Defect," Journal of the American Medical Association, 62: 741-747, Mar. 7, 1914.

Kohs, S. C., Intelligence Measurement, New York: MacMillan Co., 1923, p. 157.

Lane, Helen Schick, "Measurement of the Mental and Educational Ability of the Deaf Child," Journal of Exceptional Children, 4: 169-173, 191, May, 1938.

_____, "The Stability of Mental Test Ratings of Preschool Age Deaf Children," Oralism and Aurelism, p. 30, Feb., 1938.

MacKane, Keith, A Comparison of the Intelligence of Deaf and Hearing Children, New York: Teachers College, Columbia University, Contributions to Education, 555, 1933.

Morris, Charles M., "A Critical Analysis of Certain Performance Tests," Journal of Genetic Psychology, 54, 85-105, Mar., 1939.

Mott, A. J., "The Ninth Year of a Deaf Child's Life," American Annals of the Deaf, 44:401, 1899.

Newlee, Clara, E., "A Report of Learning Tests with Deaf Children," Volta Review, 21, 5: 216-223, Mar., 1919.

Peterson, Edwin G., "Testing Deaf Children with Koh's Block Designs," American Annals of the Deaf, 81, 3: 242-254, May, 1936.

Peterson, Edwin G., and Williams, James M., "Intelligence of Deaf Children as Measured by Drawings," American Annals of the Deaf, 75, 4: 273-293.

Pintner, Rudolf, "Contributions of Psychological Testing to the Problems of the Deaf," Proceedings of the International Congress on the Education of the Deaf, June, 1933, pp. 211-216.

_____, Intelligence Testing, New York: Henry Holt and Co., 1923, pp. 3-23, 312-327.

_____, "Latest Phases of Psychological Testing with the Deaf," American Annals of the Deaf, 82, 4: 327-337, Sept., 1937.

_____, "The Survey of Schools for the Deaf, V, Psychological Survey," American Annals of the Deaf, 72: 273-299, 377-414, 1927.

_____, "The Value of Mental Tests in the Classification of Pupils," American Annals of the Deaf, 63, 2: 196-204, Mar., 1918.

Pintner, Rudolf, and Paterson, Donald G., "A Class Test with Deaf Children," Journal of Educational Psychology, 6: 591-599, 1915.

_____, "A Combined Mental and Educational Survey," Journal of Educational Psychology, 12: 32-43, 83-91, 1921.

_____, A Scale of Performance Tests, New York: Appleton and Co., 1931 Revision.

_____, Psychological Tests of Deaf Children," Volta Review, 19, 12: 661-667, Dec., 1917.

_____, "Some Conclusions from Psychological Tests of the Deaf," Volta Review, 20: 10-14, 1918.

_____, "The Binet Scale and the Deaf Child," Journal of Educational Psychology, 6: 201, 1915.

_____, "The Survey of a Day-School for the Deaf," American Annals of the Deaf, 61, 5: 417-433, Nov., 1916.

Reamer, Jeannette Chase, "Mental and Educational Measurements of the Deaf," Psychological Monographs, 293, 1921.

Review of "Psychological Tests and Their Uses," Review of Educational Research, 8, 3: 317-317, June, 1938.

Roth, Stanley D., "Survey of the Psychological Examination Given by Dr. Stella M. Bowers, May, 1937," The West Virginia Tablet, 61, 7: 1-6, Apr., 1938.

Schick, Helen F., "A Performance Test for Deaf Children of School Age," Volta Review, pp. 657-659, Nov., 1934.

Schick, Helen F., "The Use of a Standardized Performance Test for Preschool Age Children with a Language Handicap," Proceedings of the International Congress on the Education of the Deaf, June, 1933, pp. 526-532.

Schick, H. F., and Meyer, Max F., "The Use of the Lectometer in the Testing of the Hearing and the Deaf," American Annals of the Deaf, 77, 4: 292-304, Sept., 1932.

Schieffelin, Barbara, and Schwesinger, Gladys C., Mental Tests and Heredity, New York: Galton Pub. Co., 1930, pp. 1-18, 64-69, 113, 260.

Skeels, Harold Manville, "A Study of Some Factors Influencing Performance Accomplishments of Two- and Three-year Old Children," Journal of Genetic Psychology, 40, 2: 375-398, June, 1932.

Springer, N. Norton, "A Comparative Study of the Intelligence of a Group of Deaf and Hearing Children," American Annals of the Deaf, 83, 2: 136-152, Mar., 1938.

Symposium, "Intelligence and Its Measurement," Journal of Educational Psychology, 12: 123, 195, 1921.

Terman, Lewis M., and Merrill, Maud A., Measuring Intelligence, Boston: Houghton Mifflin Co., 1937, p. 201, 319, 231.

Wells, F. L., Mental Tests and Clinical Practice, New York: World Book Co., 1927, p. 315.

Performance Scale (Ages 6 - 16)

NAME C.A.
M.A.
DATE EXAMINER I.Q.
DATE OF BIRTH

I. Block Design

	Time	Moves	Points
1.			
2.			
3.			
4.			
5.			
6.			
7.			
8.			
9.			
10.			
11.			
12.			
13.			
14.			
15.			
16.			
17.			

II. Knox Cube

1234	12434
1324	13124
1432	143124
1423	132413
13243	142341
14324	134212

III. Seguin Board

1.
2.
3.

REMARKS:

IV. Manikin

Feature profile

V. Form Boards

1. Two figure
2. Healy A
3. Casuist

VI. Picture Completion

	Points
Br. Window	
Dog	
Log	
Basket	
Cat	
Football	
Flying Bird	
Hat	
Chicken	

VII. Drawing